

Dynamical Models of Elliptical Galaxies – II. M87 and its Globular Clusters

A. Agnello^{1*}, N. W. Evans¹, A.J. Romanowsky^{2,3} and J.P. Brodie³

¹ *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

² *Department of Physics and Astronomy, San José State University, One Washington Square, San José, CA 95192, USA*

³ *University of California Observatories, 1156 High Street, Santa Cruz, CA 95064, USA*

Accepted . Received

ABSTRACT

We study the Globular Cluster (GC) system of the nearby elliptical galaxy M87 using the newly available dataset with accurate kinematics provided by Strader et al. (2011). We find evidence for three distinct sub-populations of GCs in terms of colours, kinematics and radial profiles. We show that a decomposition into three populations – blue, intermediate and red GCs – is statistically preferred to one with two or four populations. The existence of three components has been suggested before, but here we are able to identify them robustly and relate them to the stellar profile. We exploit the sub-populations to derive dynamical constraints on the mass and Dark Matter (DM) content of M87 out to ~ 100 kpc. We deploy a class of global mass-estimators, developed in Paper I, obtaining mass measurements at different locations. The DM fraction in M87 changes from ≈ 0.2 at the effective radius of the stellar light (0.02° or 6 kpc) to ≈ 0.95 at the distance probed by the most extended, blue GCs (0.47° or 135 kpc).

We supplement this analysis with *virial decompositions*, which exploit the dynamical model to produce a separation into multiple components. These yield the luminous mass as $5.5^{+1.5}_{-2.0} \times 10^{11} M_\odot$ and the dark matter within 135 kpc as $8.0^{+1.0}_{-4.0} \times 10^{12} M_\odot$. The inner DM density behaves as $\rho \sim r^{-\gamma}$ with $\gamma \approx 1.6$. This is steeper than the cosmologically preferred cusp of $\rho \sim r^{-1}$ (Dubinski & Carlberg; Navarro, Frenk & White), and may provide evidence of dark matter contraction. Finally, we combine the GC separation into three sub-populations with the Jeans equations, obtaining information on the orbital structure of the GC system. The centrally concentrated red GCs exhibit tangential anisotropy, consistent with the depletion of radial orbits by tidal shredding. The most extended blue GCs have an isotropic velocity dispersion tensor in the central parts, which becomes more tangential moving outwards, consistent with adiabatic contraction of the DM halo.

1 INTRODUCTION

Elliptical galaxies and their dark matter (DM) haloes provide crucial tests of galaxy formation scenarios. The different phases of assembly (Johansson et al. 2012; Hilz et al. 2013), the interplay between luminous and dark components over different timescales (Blumenthal et al. 1986; Abadi et al. 2010; Zubovas & King 2012), as well as the initial conditions (Stiavelli & Bertin 1987; Nipoti et al. 2012; Remus et al. 2013) all leave imprints on the final profiles. However, some properties of present-day ellipticals are not straightforward to obtain. For example, the orbital distribution of stars and the DM halo profile are not directly accessible. These must be derived by modelling the observed photometry and line-of-sight (LOS) kinematics, and are generally degenerate with one another (*mass-anisotropy degeneracy*).

Additional information on the mass profile can be gathered through the kinematics of Globular Clusters (GCs) and

Planetary Nebulae (PNe). This has been possible thanks to dedicated observational campaigns and instrumentation (such as the PN.Spectrograph, Douglas et al. 2002). Observations of PNe probe the photometric and kinematic profiles of the starlight well beyond the effective radius, where the surface-brightness would be too low (Coccato et al. 2009). Globular Cluster populations typically extend out to larger radii and, if their distribution and/or kinematics differ from those of the starlight, can be exploited as independent tracers of the same gravitational potential.

Use of GC or PN kinematics as a mass constraint dates back to Mould et al. (1987) and has been applied to a diverse set of early-type galaxies since then (e.g. Pota et al. 2013), to infer masses and properties of the DM haloes (Napolitano et al. 2009; Schuberth et al. 2010; Deason et al. 2012) or validate galaxy formation scenarios (Coccato et al. 2009; Schuberth et al. 2012). The models have relied on the Jeans equations or on distribution functions to reproduce the velocity dispersions and, in some

cases, higher-order moments of the LOS velocity distribution. Perhaps the most complete method is given by orbit-based (Schwarzschild 1979, hereafter S79) and made-to-measure modelling (Syer & Tremaine 1996; Morganti et al. 2013). We will call these *forward techniques*, to emphasize that they consist in building three-dimensional densities and velocities, which are then compared to the projected observables (surface density, velocity dispersion and higher-order moments).

However, conclusions on the dynamical side are weakened by the mass-anisotropy degeneracy, such as in the controversy over DM fractions and orbital structure (Romanowsky et al. 2003; Dekel et al. 2005). Also, the sample size and accuracy of kinematic measurements play a key role. For reference, reliable inference on the LOS velocity distribution and its higher moments requires a lower limit of $N \approx 200$ tracers per radial bin (Amorisco & Evans 2012), a condition which is not always met in practice. Especially with the limited number of GCs per host galaxy, only marginal conclusions on the orbital structure can be drawn from analyses of the higher-order moments. In the analogous context of resolved stellar populations in dwarf spheroidals (dSphs), where larger datasets are available, significant improvements have been brought by the identification of multiple populations and the use of global estimators, sometimes with sharper performance than more refined analyses in which all sub-populations were grouped together as a whole (Walker & Peñarrubia 2011; Agnello & Evans 2012b; Amorisco et al. 2013). The virial results can then be used as a prior on the dynamical parameters, to inform more detailed analyses whose outcome is mainly the orbital structure (Richardson & Fairbairn 2013).

Moreover, forward techniques require a number of steps to relate the projected distribution on the sky to the projected kinematics. Given the generally low number of GCs or PNe per galaxy (a few hundred in the best cases, so far), the surface density distribution $\Sigma(R)$ and kinematics of tracers can be very uncertain, adding to the computational complexity of some forward modelling techniques. In a companion paper (see Agnello et al. 2013b, hereafter Paper I), a novel approach has been illustrated. This is based upon a reformulation of the Jeans equations, which yields the predicted velocity dispersions and their aperture-averages by means of single or double integrals involving $\Sigma(R)$ directly, along with a kernel that depends on the mass and anisotropy profiles. As a by-product, different mass estimators at particular locations have been derived, as well as characterisations of the kinematic profiles and common aperture-corrections (cf Jorgensen et al. 1995; Cappellari et al. 2006). Some particular cases (the virial limit and isotropic tracers in power-law potentials) have been previously exploited in a variety of problems, ranging from gravitational lensing by galaxies (Agnello et al. 2013) to the DM profiles of dwarf spheroidals (Amorisco et al. 2013).

Here, we will apply the results of Paper I to the early-type galaxy M87 and its GC system. This is one of the most luminous and massive galaxies within 20 Mpc, and hosts one of the most populous GC systems. It was the first target of GC spectroscopy beyond the Local Group (Mould et al. 1987, 1990; Huchra & Brodie 1987). Since then, its GCs have been studied extensively, most notably with larger kinematic samples by Cohen & Ryzhov (1997), Cohen (2000)

and Hanes et al. (2001). Interest then turned to other galaxies until Strader et al. (2011, hereafter S11) revisited M87 with a new generation of high resolution, high signal-to-noise spectroscopy, providing the largest, publicly available, sample in the literature. This is the dataset we will use here.

In Section 2, we introduce the dataset that will be used in the subsequent dynamical analyses and summarize previous results in the literature. In Section 3, we present a partition into sub-populations, showing that the GC system of M87 is most likely a superposition of three distinct (and internally homogeneous) sub-populations. In Section 4, we use the three GC populations to obtain virial constraints on the mass profile (luminous mass, DM fractions, DM contraction) in a manner that is independent of any orbital structure. In Section 5, we then apply the integral form of the Jeans equations derived in Paper I to learn about the orbital distributions. We comment and draw conclusions in Section 6.

From now on, we adopt a systemic velocity of 1307 km s^{-1} and a distance of 16.5 Mpc for M87 (see e.g., Madrid et al. 2009; Bird et al. 2010). Throughout all this paper, fits and likelihood explorations are performed using the Markov Chain Monte Carlo technique.

2 M87 AND ITS GLOBULAR CLUSTERS

The S11 data were based on wide-field imaging from¹ Subaru/Suprime-Cam and CFHT/MegaCam, with follow-up spectroscopy on Keck/DEIMOS and MMT/Hectospec out to 185 kpc (≈ 0.64 degrees). These data have undergone extensive vetting for reliability and the velocities are accurate at the $\approx 10 \text{ km s}^{-1}$ level. Comparison to the older generation of data (performed by S11) was however revealing. Several cases of ‘catastrophic’ errors were found, with differences up to $\approx 10^3 \text{ km/s}$ that seem most likely to be problems with the older data on a smaller telescope. Unfortunately, even a few such unrecognized errors can have dramatic impacts on kinematic parameters such as the velocity dispersion, with subsequent repercussions on the dynamical mass. In particular, the GC dispersion profile of M87 was previously thought to rise rapidly with radius, but with the newer data it is seen to remain constant or decline (as shown below and in S11).

The kinematics of smaller datasets of GCs at smaller radii have already been studied before. Cohen & Ryzhov (1997) performed a Jeans analysis, assuming an isotropic velocity dispersion tensor, on 205 GCs between ≈ 0.03 and 0.14 degrees (i.e. $10 \lesssim R/\text{kpc} \lesssim 50$). They used a power-law model for the total density (i.e. luminous plus DM),

$$\rho_{\text{tot}}(r) \propto r^{-\gamma}, \quad (1)$$

and inferred $\gamma \approx 1.3$, corresponding to a (total) mass profile $M(r) \propto r^{1.7}$. Romanowsky & Kochanek (2001) studied a set of 200 GCs by means of S79 orbit modelling, finding evidence for a dark halo with density falling off more slowly than r^{-2} . Wu & Tremaine (2006) applied a S79 orbit-based analysis,

¹ Acronyms: CHFT, Canada-France-Hawaii Telescope; DEIMOS, DEep Imaging Multi-Object Spectrograph; MMT, Multiple Mirror Telescope. See S11 and references therein.

with a power-law total density, to a sample of 161 GCs between ≈ 0.03 and 0.3 degrees. They found $\gamma = 1.6 \pm 0.4$, although the likelihood has three peaks, one of which is almost at $\gamma = 2$. The energy distribution resulting from their modelling suggested multiple distinct components, although the data were not enough to confirm this hypothesis robustly. Murphy et al. (2011) analysed 278 GCs up to 47 kpc (≈ 0.15 deg) by means of orbit-based methods, together with stellar kinematics for the starlight, as to infer the (total) mass profile over a wide radial range. The expected (total) mass at 2 32 kpc is $\approx 34\%$ higher than the one reported in Wu & Tremaine (2006). Also, mass-estimates from X-rays are almost 50% lower than the ones inferred from GCs at small radii (Murphy et al. 2011).

Here, we make use of the S11 dataset, while omitting the brightest objects (with magnitude $i < 20$) that may be a distinct population of ultra-compact dwarfs (as discussed in S11 and Brodie et al. 2011). We also do not make use of the additional DEIMOS data from Romanowsky et al. (2012), since these were specifically obtained around a cold substructure that could bias the kinematics. In particular, the sample with accurate kinematic measurements (the ‘kinematic’ sample) is a subset of the dataset studied in S11 (the ‘photometric’ sample).

2.1 The Data

In order to exclude possible contaminants from our catalogue, we adopt the same criterion of Harris (2009) and S11, considering just those objects in the kinematic sample with colours as follows:

$$0.43 < g-r < 0.95, \quad 0.59 < g-i < 1.40. \quad (2)$$

Our kinematic sub-sample consists of 420 bona fide GCs, with positions, LOS velocities and colours. A clear bimodality is present in the colour distribution (as already observed in Harris 2009, S11).

A priori, the selection function for the kinematic sub-sample may not be the same as for the photometric catalogue, both in positions and in colours. If this happened, the surface-density profiles $\Sigma(R)$ used in the dynamical analysis would be inadequate and any mass measurements would be affected by that. Also, we must be confident that just the *bona fide* GCs are considered for the dynamical analysis. Harris (2009) and S11 examined different procedures to isolate the GCs, subtract the contaminants and separate the sample in bluer ($g-i < 0.93$) and redder ($g-i > 0.93$) sub-populations.

Figure 1 shows the estimated radial profiles, for bluer and redder GCs. The profiles given in S11 are consistent with the number counts for pure GCs estimated by Harris (2009) with a different procedure for contaminant subtraction. The surface-density profile of ‘red’ GCs matches smoothly to the starlight’s surface-brightness profile in the inner regions, suggesting that the starlight is associated with a red GC population. This fades at larger radii into a bluer, metal-poorer and less concentrated component, which may be the result of accretion onto M87 (Brodie & Strader 2006).

² Or at 35 kpc with their adopted distance, which is 17.9 Mpc, whereas Wu & Tremaine (2006) assumed ≈ 16 Mpc.

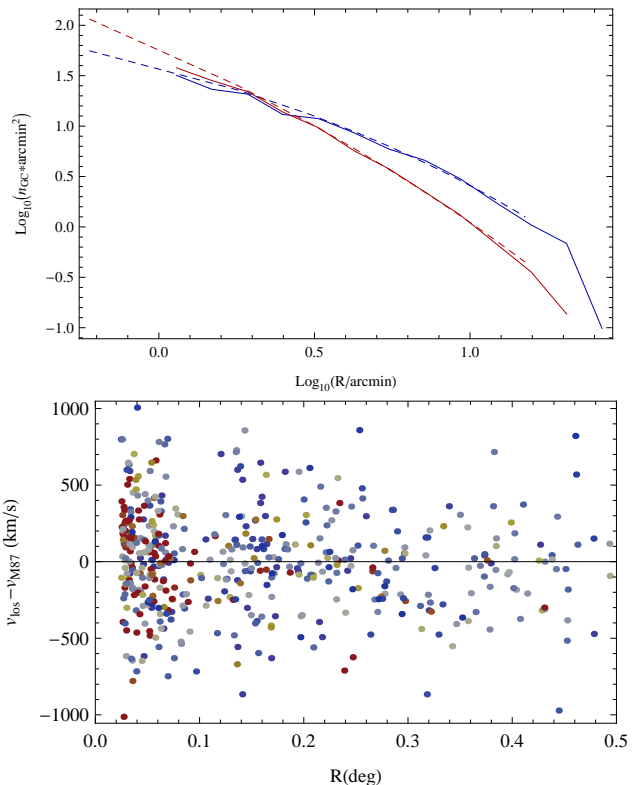


Figure 1. M87 Globular Cluster data.

Top: Surface density profiles of GCs in M87, after subtraction of contaminants. The full lines are given by the number-counts of Harris (2009), where the surface-density of contaminants is estimated directly from field objects; the dashed lines show the best-fitting profiles of S11, where an independent photometric sample was used, the fraction of contaminants was a free parameter in the modelling and the radial selection function was accounted for. Blue (dark red) denotes the GCs with $g-i$ lower (higher) than 0.93 . Note that the redder GCs are more centrally concentrated. Bottom: Line of sight velocities (relative to M87) versus distance from M87 for GCs in the spectroscopic subsample of S11. Redder (bluer) points mark GCs with higher (lower) values of $g-i$.

In the dynamical analysis, we will use data from the kinematic sub-sample. Then, we need to know if the photometric profiles are consistent between the photometric sample and the kinematic subsample. Although the radial selection function is different, a simple check consists in examining the ratios of redder-to-bluer GCs at different locations, which depends just on the colour selection function. Figure 2 shows the ratio of redder-to-bluer GCs estimated by ratio of number-counts in different radial bins, both for the photometric and kinematic sub-samples. As the plots show, the kinematic sub-sample is a faithful representation of the total (photometric) sample without substantial colour biases or contamination in the distance range probed here.

3 SUB-POPULATIONS

A simple colour separation based on hard cuts ³ shows that the bluer GCs have a higher velocity dispersion than the

³ That is, $g-i$ larger or smaller than the conventional value 0.93 .

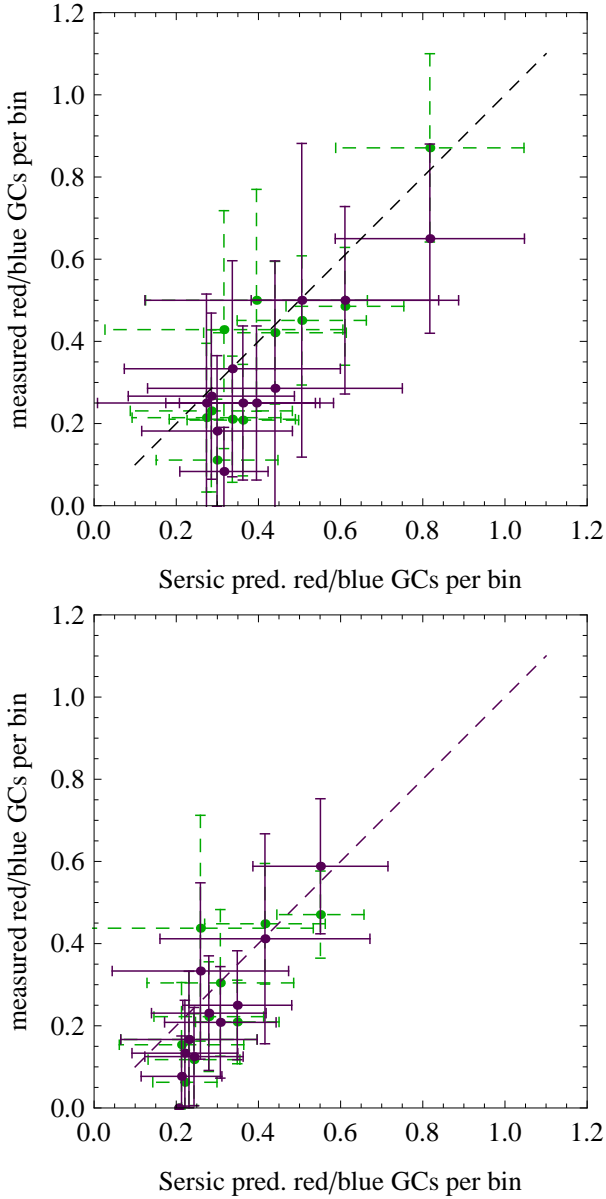


Figure 2. The selection function for the photometric and kinematic samples, see Section 2.1. Green dashed (purple solid) error bars denote the photometric (kinematic) sample. The upper panel uses bin sizes of 0.02° , the lower panel 0.04° . The dashed line shows the 1:1 proportionality. The selection function has no appreciable variation with colour and has the same radial profile for both the photometric and spectroscopic datasets.

redder ones. Having both colours and LOS velocities, we can exploit them jointly to separate our sample into sub-populations.

We will suppose that each population has a Gaussian distribution in colour and velocity, of the kind

$$\mathcal{G}(v, c; \sigma_j, \langle c \rangle_j, \Delta c_j) = \frac{1}{2\pi\sigma_j\Delta c_j} \exp \left[-\frac{v^2}{2\sigma_j^2} - \frac{(c - \langle c \rangle_j)^2}{2\Delta c_j^2} \right], \quad (3)$$

(c.f., Walker & Penarrubia 2011). We use the difference between g and i magnitudes as the colour $c = g - i$ in our study, while j denotes the population, v the line-of-sight velocity,

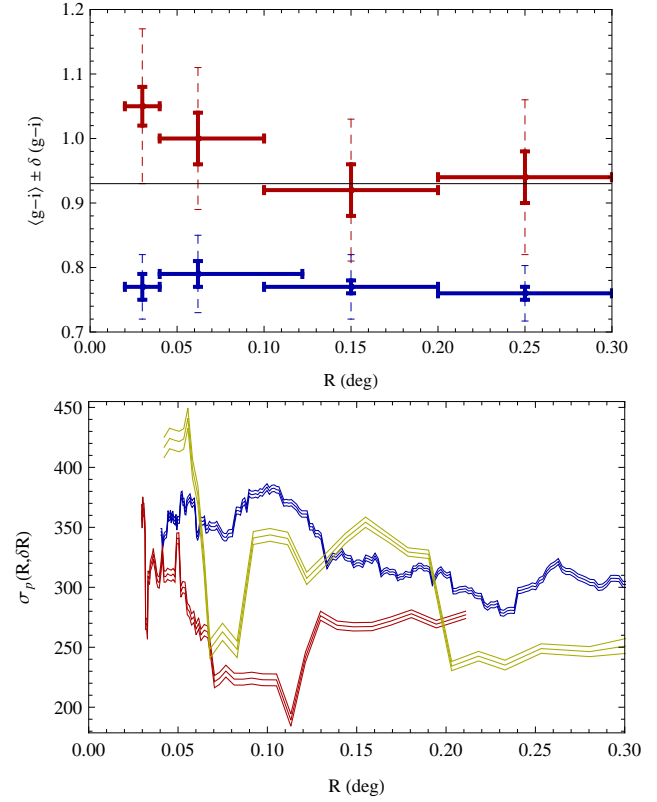


Figure 3. Behaviour of GC colours and kinematics with distance from M87.

Top: Colour $(g - i)$ parameters resulting from a bimodal fit in different bins of distance (horizontal error-bars), as in Sect.3.1 and Table 1. Full error-bars: mean colour and its uncertainty; dashed errorbars: width of the colour distribution, centered on the mean colour for visual convenience. The mean-colour variation of the ‘red’ GCs over distance is monotonic, appreciably larger than the uncertainties and is comparable to the colour dispersion, at variance with the behaviour of ‘blue’ GCs.

Bottom: Velocity dispersion profiles for GCs of different colours. Blue (yellow, red): GCs with $g - i < 0.93$ ($0.93 < g - i < 1$, $g - i > 1$). Full lines mark the velocity dispersions computed via eq.(5), whilst dashed lines delimit the uncertainties computed by linear propagation on v_k in eq.(5). The bin-size δR varies with radius, such that a quarter of the GC subpopulation is enclosed in each bin. GCs with $g - i \geq 0.93$ show a sharp transition in $\sigma_p(R, \delta R)$, which corresponds to the transition in mean colour of the redder GCs (top panel).

σ the velocity dispersion and $\langle c \rangle, \Delta c$ the mean colour and colour dispersion. The above distribution can be modified to distinguish between rotation and random motions. However, in what follows, we are primarily interested in the entirety of the second velocity moment, whether due to ordered or random motions, and this is the same as the dispersion σ_j estimated via eq. (3).

3.1 Bimodal Partition

We start by dividing our sample into four radial bins and performing a bimodal fit in each of them. We suppose that $(\sigma, \langle c \rangle, \Delta c)$ are reasonably uniform for each sub-population in each radial bin. In the bimodal fitting, each GC with

bin (deg)	N_{bin}	$\langle c \rangle_{\text{r}}$	$\langle c \rangle_{\text{b}}$	Δc_{r}	Δc_{b}	σ_{r} (kms $^{-1}$)	σ_{b} (kms $^{-1}$)	f_{b}
$0.02 < R < 0.04$	85	1.05 ± 0.03	0.77 ± 0.02	0.12 ± 0.02	0.05 ± 0.01	330 ± 37	430 ± 65	0.38 ± 0.09
$0.04 < R < 0.1$	123	1.01 ± 0.04	0.79 ± 0.02	0.11 ± 0.02	0.06 ± 0.01	265 ± 29	374 ± 44	0.51 ± 0.12
$0.1 < R < 0.2$	78	0.92 ± 0.04	0.77 ± 0.01	0.11 ± 0.02	0.05 ± 0.01	280 ± 56	315 ± 57	0.54 ± 0.11
$0.2 < R < 0.3$	68	0.94 ± 0.04	0.76 ± 0.01	0.12 ± 0.02	0.04 ± 0.01	295 ± 48	290 ± 40	0.59 ± 0.10

Table 1. Colour parameters, velocity dispersions and blue to red fractions in different radial bins from a bimodal fit to the colour and velocity distribution. Here, subscripts ‘r’ and ‘b’ refer to the redder and bluer populations respectively. The errorbars on $\langle g-i \rangle$ are significantly smaller than the colour variation with radial bin, suggesting a real change of mean colour (and metallicity) with radius in the redder component. The column N_{bin} indicates the number of GCs in each radial range.

measured velocity $v_k \pm \delta v_k$ and colour $c_k \pm \delta c_k$ gives a contribution

$$\mathcal{L}_k = f_{\text{b}} \mathcal{G} \left(v_k, c_k; \sqrt{\sigma_{\text{b}}^2 + \delta v_k^2}, \langle c \rangle_{\text{b}}, \sqrt{\Delta c_{\text{b}}^2 + \delta c_k^2} \right) + (1 - f_{\text{b}}) \mathcal{G} \left(v_k, c_k; \sqrt{\sigma_{\text{r}}^2 + \delta v_k^2}, \langle c \rangle_{\text{r}}, \sqrt{\Delta c_{\text{r}}^2 + \delta c_k^2} \right) \quad (4)$$

to the likelihood of the bimodal partition, with f_{b} (or $(1 - f_{\text{b}})$) the fraction of ‘bluer’ (or ‘redder’) GCs in the radial bin considered. The final likelihood for the $(\sigma, \langle c \rangle, \Delta c)$ parameters is the product of individual contributions $\mathcal{L} = \prod_k \mathcal{L}_k$ over all objects in the bin. Different possible bin sizes were investigated, with the aim of minimising uncertainties and bin size while keeping the number of GCs equal among different bins.

Table 1 summarises the outcome of our bimodal fitting and the bin sizes chosen. The blue fraction f_{b} increases with radius, as expected based on Harris (2009) and S11. The ‘bluer’ sub-population has overall constant colour parameters and a smoothly declining velocity dispersion profile. The ‘redder’ GCs have a mean colour that decreases towards $\langle g-i \rangle \approx 0.93$. Also, their LOS kinematics shows an abrupt drop-off from $\approx 330 \text{ kms}^{-1}$ in the innermost bin to $\approx 270 \text{ kms}^{-1}$ at larger radii. This is accompanied by a significant change in the mean colour of the ‘redder’ sub-population. The inner value of the velocity dispersion is consistent with the kinematics of the stellar component (S11). Suppose the velocity dispersion of ‘redder’ GCs is estimated at different radii via a uniform smoothing kernel

$$\sigma_{\text{p}}^2(R; \delta r) = \frac{1}{N} \sum_{k=1}^N v_k^2 \mathbf{1}_{|R_k - R| < \delta r}, \quad (5)$$

where $\mathbf{1}$ is the Heaviside function. Then, the dispersion still shows a sharp decrease for the ‘redder’ GCs, which does not have a counterpart in the ‘bluer’ ones (fig.3). On the other hand, the LOS kinematic profile for the starlight is approximately flat even beyond the effective radius, so we would not expect a sharp drop in $\sigma_{\text{p}}(R)$ for the red GCs either.

The following hypothesis emerges: the red GCs consist of a more compact (and redder) sub-population and a more extended, intermediate-colour sub-population. This hypothesis was suggested and supported also in Sections 5 and 6 of S11. Our sample is a subset⁴ of the one considered by S11,

⁴ Besides those used here, S11 included other GCs, with mea-

parameter	blue	intermediate	red
$\langle c \rangle$	0.773 ± 0.005	0.93 ± 0.03	1.10 ± 0.02
Δc	0.047 ± 0.004	0.09 ± 0.01	0.074 ± 0.009
σ (km/s)	350 ± 19	275 ± 25	333 ± 37
n	$2.25^{+1.45}_{-0.45}$	$4.0^{+1.1}_{-1.8}$	$4.68^{+0.90}_{-1.10}$
$R_{\text{e}}(\text{arcsec})$	$\Sigma_{\text{i/b}} = 20^{+10}_{-7}, \Sigma_{\text{r/b}} = 2480^{+3120}_{-1060}$ 2300 ± 700	1750 ± 650	72^{+128}_{-22}
$\langle g-i \rangle$	0.773 ± 0.004	0.92 ± 0.02	1.09 ± 0.03
Δc	0.047 ± 0.003	0.09 ± 0.02	0.08 ± 0.02
σ (km/s)	350 ± 15	267 ± 12	326 ± 34
$f_{\text{b}} = 0.54 \pm 0.05, f_{\text{i}} = 0.34^{+0.07}_{-0.13}$			

Table 2. The dissection into three populations with (upper) and without (lower) positional information. In the upper table, we give colour parameters, mean velocity dispersions and Sérsic indices for the three sub-populations. The remaining photometric parameters are not well constrained, so underneath we list the effective radii and normalisations once $\langle g-i \rangle_j$, δc_j , σ_j and n_j are fixed to their maximum likelihood values. The posterior on effective radii has a large width (cf S11), especially for the blue and intermediate populations. In the lower table, we list the parameters of a separation into three components in which no information from the positions is used. Reassuringly, the best-fit colour parameters and velocity dispersions do not change significantly. Also, the parameters $R_{\text{e,r}}, R_{\text{e,i}}, \Sigma_{\text{i/b}}/\Sigma_{\text{r/b}}$ are very close to the ones derived by Janowiecki et al. (2010) on the starlight.

which had been thoroughly examined to exclude contaminants or evident substructure and select just the GCs that are bound to M87. Here, we will make one further step and exploit this hypothesis to identify three dynamical tracer populations, which can be used to infer masses at different locations.

3.2 Three Populations

We proceed to a decomposition into three sub-populations, each of which has uniform colour parameters and velocity

measurements already known from the literature but with unverified velocities.

dispersion. We model each sub-population with a Sérsic profile for the surface number-density

$$\Sigma_j(R) = \Sigma_{0,j} \exp \left[-\kappa_{n_j} (R/R_{e,j})^{1/n_j} \right], \quad (6)$$

and denote the fraction of j -th population at radius R as

$$f_j(R) = \frac{\Sigma_j(R)}{\Sigma_b(R) + \Sigma_r(R) + \Sigma_i(R)}, \quad (7)$$

where we have added a new subscript ‘i’ to denote the intermediate population. The Sérsic coefficient κ_n is chosen such that $R_{e,j}$ is the effective radius of the j -th population, enclosing half the total (projected) number of GCs. In particular, we rely on the approximate expression listed in Ciotti & Bertin (1999) in terms of the Sérsic index n .

Then, a GC at radius R_k will give a contribution to the likelihood

$$\begin{aligned} \mathcal{L}_k = & f_b(R_k) \mathcal{G} \left(v_k, c_k; \sqrt{\sigma_b^2 + \delta v_k^2}, \langle c \rangle_b, \sqrt{\Delta c_b^2 + \delta c_k^2} \right) \\ & + f_r(R_k) \mathcal{G} \left(v_k, c_k; \sqrt{\sigma_r^2 + \delta v_k^2}, \langle c \rangle_r, \sqrt{\Delta c_r^2 + \delta c_k^2} \right) \\ & + f_i(R_k) \mathcal{G} \left(v_k, c_k; \sqrt{\sigma_i^2 + \delta v_k^2}, \langle c \rangle_i, \sqrt{\Delta c_i^2 + \delta c_k^2} \right). \end{aligned} \quad (8)$$

Just as for the bimodal fits, the likelihood of the parameters is the product of \mathcal{L}_k over all GCs in our sample. The likelihood is now a function of: effective radii R_e and Sérsic indices n of the populations; mean colours and colour dispersions; global velocity dispersions; and the ratios $\Sigma_{r/b} = \Sigma_r(0)/\Sigma_b(0)$, $\Sigma_{i/b} = \Sigma_i(0)/\Sigma_b(0)$ between the surface number densities at the centre. By studying the likelihood over the parameter space, we obtain both the best-fitting parameters and their distribution, including uncertainties.

Table 2 shows the colour parameters, mean velocity dispersions and photometric parameters. Uncertainties on the photometric parameters are large, especially for the effective radii. This was also observed in S11, where photometric profiles were sought for the bluer ($g-i < 0.93$) and redder ($g-i > 0.93$) GCs in the larger, photometric sample. Despite the uncertainties, some well-defined relations are satisfied among the parameters. In particular, the effective radius of the intermediate population is almost always smaller than that of the blue GCs (with $P(R_{e,i} > R_{e,b}) < 10\%$). Also, the velocity dispersions are strictly ordered as $\sigma_b > \sigma_r > \sigma_i$ ($P(\sigma_i > \sigma_r) < 10\%$ and $P(\sigma_r > \sigma_b) < 4\%$).

We can also examine a fit with three components in which no information on the radial positions is used. This is a way of checking that the findings on colours and kinematics are robust. In this case, all GCs are grouped together in the same distance bin and the fractions f_b , f_i , $f_r = 1 - f_b - f_i$ are free parameters themselves. The result is shown in the bottom section of Table 2 and confirms the findings based on radial profiles.

We also investigate a different choice of velocity dispersion law, which can be rising or falling at large radii (cf Table 1), namely⁵

$$\sigma_p(R) = \sigma_0 + \frac{\sigma_1 R_a^\zeta}{(R_a^2 + R^2)^{\zeta/2}}. \quad (9)$$

⁵ With a spatially varying velocity dispersion, the likelihood is simply modified by replacing σ_j with $\sigma_j(R_k)$ for the k -th GC.

number of populations	$\log \mathcal{L}$	$\log_{10} Z$
2	-2694.0	-1177.4
3	-2691.8	-1175.4
4	-2691.0	-1175.2

Table 3. Partitions using information on colours and velocities, but not on distances from the center of M87. The maximum likelihood of each decomposition is listed ($\log \mathcal{L}$) as well as the logarithmic evidence ($\log_{10} Z$).

Use of equation (9) for the kinematics of any of the subpopulations does not improve the likelihood of the fit appreciably, as the increase in $\log \mathcal{L}$ is not sufficient to balance the increase in degrees of freedom. Moreover, the global velocity dispersion of each population $\sqrt{\langle \sigma_{p,j}^2 \rangle}$, computed from

$$\langle \sigma_p^2 \rangle = \frac{\int_0^\infty R \Sigma(R) \langle \sigma_{p,j}^2 \rangle (R) dR}{\int_0^\infty R \Sigma(R) dR}, \quad (10)$$

does not differ appreciably from the result of a fit with uniform σ . In other words, the fit deploying a parameterization of σ yields some marginal information on the velocity dispersion profiles, without changing the global velocity dispersion of each population. The effective radii $R_{e,r}$, $R_{e,i}$ and the surface-density ratio $\Sigma_{i/b}/\Sigma_{r/b} = \Sigma_i(0)/\Sigma_r(0)$ are remarkably close to the findings of a double de Vaucouleurs fit to the starlight, as performed by Janowiecki et al. (2010). This lends further support to the robustness of our results and suggests a link between GC populations and the stellar component.

An important fact is the independence of these findings from any dynamical model for M87 and its gravitational potential. Then, for a dynamical model to be reliable, it must be able to reproduce the main features of this decomposition. A criterion to assess the acceptability of a dynamical model will be whether the ordering $\sigma_b > \sigma_i > \sigma_r$ in velocity dispersions and $R_{e,b} > R_{e,i}$ in effective radii are satisfied.

3.3 Statistical Evidence for Three Populations

The separation into three components gives a smoother description of the GC system of M87. To ensure that this choice is indeed preferable, we must quantify how well it describes the colours and kinematics with respect to a partition into two sub-populations. We consider two different *Bayes factors* (Jeffreys 1961; Kendall & Stuart 1979; Burnham & Anderson 2002) as quantitative criteria, namely the ratio of maximum likelihoods and of evidences.

Given two models A and B with maximum likelihoods $\mathcal{L}_A > \mathcal{L}_B$, let us define

$$\Delta\chi^2 = 2 \log \mathcal{L}_A - 2 \log \mathcal{L}_B. \quad (11)$$

In the case of least-square fitting with Gaussian statistics, this would be exactly the decrease in χ^2 . Now let the number of free parameters in A exceed the one in B by an amount Δp . Then, the larger the difference

$$B_1 = \Delta\chi^2 - \Delta p, \quad (12)$$

pop.	R_* (deg)	$M(R_*)$ (M_\odot)	$v_c(R_*)$ (km/s)	M/L (M/L) $_{\odot,V}$	f_d
red	0.01	1.7×10^{11}	500	17.0	0.1
int.	0.10	1.2×10^{12}	420	31.6	0.55
blue	0.47	9.0×10^{12}	530	129.0	0.94
starlight	0.02	3.4×10^{11}	500	22.7	0.2

Table 4. Masses enclosed at the pinch radii of the blue, intermediate and red populations, together with implied circular velocities. The quoted values of R_* correspond to the peak of the posterior distribution in logarithmic bins. The quoted value of mass corresponds to the peak of the distribution in $M(R_*)$, inherited from the distribution of σ_j , $\log R_{e,j}$ and n_j via eq. (17). The V-band mass-to-light ratio is computed adopting the double de Vaucouleurs fit to the starlight described by Janowiecki et al. (2010). The final column gives the dark matter fraction, using the luminosity profile discussed in McLaughlin (1999), although comparable results are obtained with different parameterisations (such as in Romanowsky & Kochanek 2001).

the more model *A* is preferable to model *B*. Equation (12) then defines the first Bayes factor, B_1 .

On the other hand, by introducing a new set of parameters, there is a risk of over-fitting. To avoid a fine-tuning of the new parameters to reproduce the observations, it is useful to examine the evidence ratio

$$\frac{Z_A}{Z_B} \equiv \frac{P(A|\text{data})}{P(B|\text{data})} = \frac{\int \mathcal{L}(\mathbf{p}_A)\theta(\mathbf{p}_A|M_A)d\mathbf{p}_A}{\int \mathcal{L}(\mathbf{p}_B)\theta(\mathbf{p}_B|M_B)d\mathbf{p}_B}, \quad (13)$$

where θ indicates the priors and \mathcal{L} the likelihoods in the two models M_A, M_B . In particular, we will quote the results in terms of

$$B_2 = \Delta \log_{10} Z = \log_{10}(P(A|\text{data})/P(B|\text{data})) \quad (14)$$

as the second Bayes factor.

Table 3 shows the results for a decomposition into two, three or four sub-populations, where no information from the positions of the GCs is used. Each additional population increases the number of free parameters by 4 ($f_i, \sigma_i, \Delta c_i$ and $\langle c \rangle_i$). On moving from two populations to three, $\Delta\chi^2 = 2\Delta\mathcal{L} = 4.4$ so that the first Bayes factor B_1 indeed indicates that a separation into three populations is preferred over two. On moving from three populations to four, $\Delta\chi^2 = 2\Delta\mathcal{L} = 1.6$, which does not outweigh the increase in the numbers of free parameters. This can be confirmed by examining the evidence, using the second Bayes factor B_2 . On comparing the hypothesis of three populations against two, $B_2 = 2.0$ so that the evidence in favour of three populations is ‘very strong’ on the Jeffreys (1961) scale. By contrast, the evidence in favour of four populations over three is ‘barely worth mentioning’ ($B_2 = 0.1$).

Another way of testing if three populations are actually a better fit is to consider the gain in the first Bayes factor arising when information on the radial positions is included. Now the two population likelihood is $\mathcal{L} = -2673$ whilst the three population likelihood is $\mathcal{L} = -2662$. With an increase $\Delta p = 6$ (a Sérsic index and scalelength in addition to the previous listed 4 parameters) and a gain $\Delta\chi^2 = 2\Delta\mathcal{L} = 22$, we can be confident that the GC system of M87 is indeed the superposition of three sub-populations.

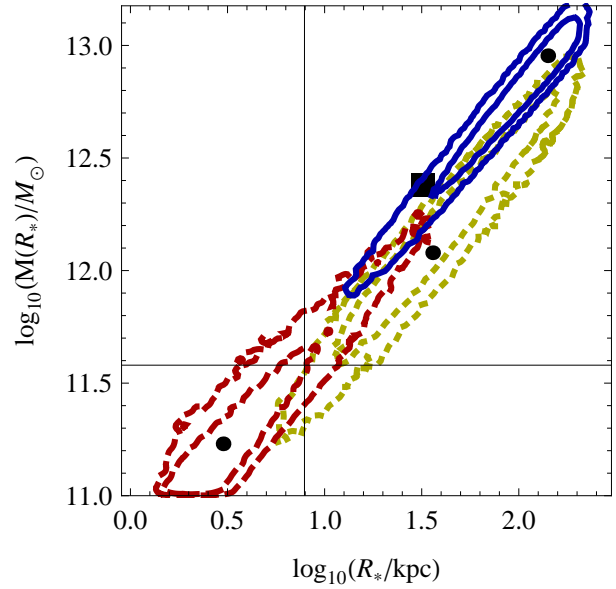


Figure 4. Distribution of the virial pinch radius R_* and of the (total) mass enclosed within it, using eq.s (16) and (17). The distribution of R_* and $M(R_*)$ is obtained from the posterior on the parameters in the three-population dissection. Blue and solid (yellow and dotted, red and dashed respectively) contours are the 68% and 95% confidence regions as given by the blue (intermediate, red) population. The axes intersect at the value estimated from the starlight. A square marks the value found by Wu & Tremaine (2006).

4 DYNAMICS

Multiple stellar populations can provide a powerful constraint on the underlying gravity field. This is because each population must be in dynamical equilibrium in the same gravitational potential. A number of authors have exploited this fact to constrain the mass distribution, particularly in dwarf spheroidal galaxies (e.g., Walker & Penarrubia 2011, Amorisco & Evans 2012). For the sake of comparison, the analyses of multiple populations in dwarf spheroidals could rely on datasets with more than 10^3 stars each, whereas here we have just 420 confirmed GCs at our disposal.

4.1 Masses within Pinch Radii

Before embarking on elaborate models, we begin with something simple. For any stellar distribution with velocity dispersion σ and effective radius R_e , Paper I shows how to identify a pinch radius R_* , within which the uncertainty in the (total) enclosed mass is minimised. We summarise the method here for the reader’s convenience.

When the kinematics of a tracer population are probed over large radii⁶, the mass within the pinch radius is:

$$M(R_*) = K \frac{R_* \langle \sigma^2 \rangle}{G}, \quad K \approx 2.3. \quad (15)$$

⁶ The necessary radial coverage is better discussed in Paper I. Here it suffices to know that the necessary requirements for these virial analyses are met. This also results a posteriori from the Jeans analysis of Section 4.4, which essentially adds further insight on the GC orbits but not on the mass profile.

model	$\hat{\sigma}_b$ (km/s)	$\hat{\sigma}_i$ (km/s)	$\hat{\sigma}_r$ (km/s)
broken power law (eq 22)	332 ± 15	318 ± 16	296 ± 25
luminous & dark (eq 23)	345 ± 20	290 ± 22	316 ± 35

Table 5. Inferred velocity dispersions, obtained by marginalising the posterior \mathcal{L} over the photometric parameters and normalisations $\Sigma_{r/b}$, $\Sigma_{i/b}$ of central surface densities.

The proportionality coefficient $K \approx 2.3$ is an average over different mass models. For a Sérsic profile of effective radius R_e and index n , the pinch radius is

$$R_* = R_e \kappa_n^{-n} \sqrt{2\Gamma(3n)/\Gamma(n)}. \quad (16)$$

The relation (15) is more conveniently restated in our case as

$$\begin{aligned} M(R_{*,j}) &= 1.54 \left(\frac{\sigma_j}{100 \text{ km/s}} \right)^2 \left(\frac{R_{*,j}}{\text{deg}} \right) \times 10^{12} M_\odot \\ &= 0.53 \left(\frac{\sigma_j}{100 \text{ km/s}} \right)^2 \left(\frac{R_{*,j}}{10 \text{ kpc}} \right) \times 10^{11} M_\odot \end{aligned} \quad (17)$$

for the j -th population (i.e. blue, red or intermediate-colour). For multiple populations, such formulae can be applied to each population, giving insights into the variation of the mass with radius. In our case, the values of R_j , σ_j , n_j are given by the MCMC likelihood exploration in the partition into sub-populations.

The results for the three sub-populations are given in Table 4 and visualised in Fig. 4 (see Walker & Peñarrubia 2011; Amorisco et al. 2013, for a similar approach). We also quote the pinch radius and enclosed mass for the starlight, simply described by a de Vaucouleurs profile with effective radius $R_e = 0.02$ degrees (Harris 2009) and velocity dispersion $\sigma_p \approx 330$ km/s (see S11 and references therein). These simple estimates suffer from the systematics that we have illustrated in Paper I. In particular, they do not account for the distribution of K given by the (unknown) mass profile. Notice that the likelihood contours in Fig. 4 are distended along the line $M(R) \propto R$, which corresponds to isothermal. This is because the relative uncertainties in the effective radii are larger than those in σ_j , n_j and the use of formulae akin to eq. (15, 16) causes this to propagate linearly into the uncertainty on the enclosed mass. Barring systematic uncertainties (see Paper I), the total mass enclosed within the most likely pinch radius has a relative uncertainty of ≈ 0.1 dex for each population.

4.2 Scale-free Total Density

Given a model $\rho_{\text{tot}}(r)$ for the total density, all population must satisfy the projected virial theorem simultaneously

$$\langle \sigma_p^2 \rangle_j = \frac{16\pi G}{3L_{\text{tot},j}} \int_0^\infty y \Sigma_j(y) \int_0^y \frac{r^2 \rho_{\text{tot}}(r) dr}{\sqrt{y^2 - r^2}} dy \quad (18)$$

(see e.g., Agnello & Evans 2012a; Amorisco et al. 2013, and Paper I). Here, $\langle \sigma_p^2 \rangle_j$ is the average velocity second moment and $\Sigma_j(R)$ is the surface density profile of the j -th population, whilst $L_{\text{tot},j}$ is the total luminosity of the population $2\pi \int_0^\infty R \Sigma_j(R) dR$.

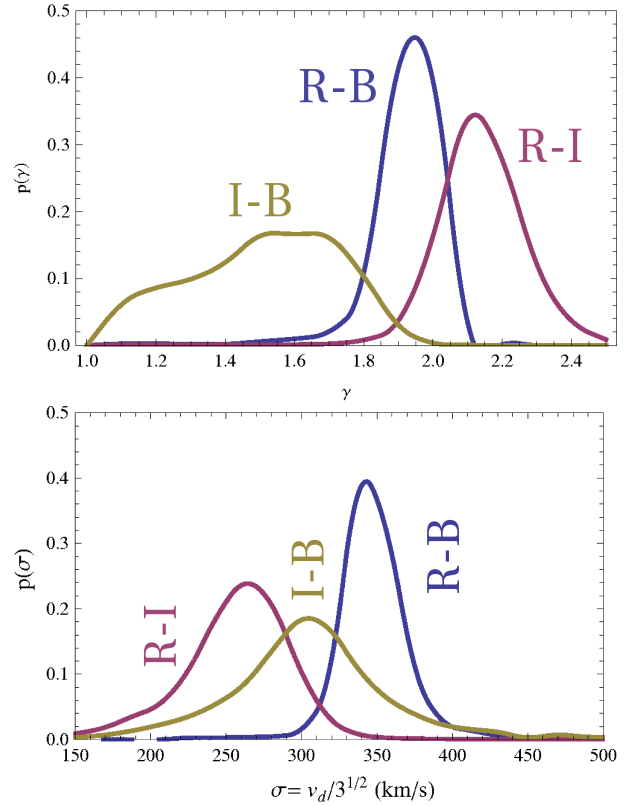


Figure 5. Marginalised likelihood for the density exponent (top) and velocity normalisation (bottom, cf eq.20) for power-law models, using the red and blue (blue curve, marked R-B), red and intermediate (purple, R-I) or intermediate and blue (yellow, I-B) sub-populations.

As a first pass, let us begin with a scale-free total density. Although an over-simplification, the scale-free approximation has often led to useful insights in the past (c.f., Churazov et al. 2010; Agnello & Evans 2012a). The total density is given by

$$\rho_{\text{tot}}(r) = \rho_0 (r/r_0)^{-\gamma}. \quad (19)$$

This has two independent quantities, the exponent γ and a normalisation ρ_0 . Given a reference radius R_d , it is convenient to define the velocity parameter

$$v_d = v_c(R_d) = \sqrt{\frac{4\pi G \rho_0 R_d^{2-\gamma} r_0^\gamma}{3-\gamma}}, \quad (20)$$

which is the circular velocity at the radius R_d . When $\gamma = 2$ (flat rotation curve), the dependence on R_d vanishes and the average velocity dispersion for all sub-populations is simply $v_d/\sqrt{3}$, as can be deduced from eq. (18). The scale R_d can be chosen arbitrarily. It just corresponds to the reference radius for mass measurements. Given our data, we have used $R_d = 10^3$ arcseconds for convenience.

From the virial theorem applied to the j -th population, we obtain a predicted velocity dispersion $\hat{\sigma}_j(v_d, \gamma)$, averaged over the population's profile as in eq. (18). We show its dependence on the two model parameters explicitly, but it also depends on the two photometric parameters ($R_{e,j}$, n_j). These, in turn, come from the partition into three populations and every choice of the photometric parameters is

weighted with the likelihood \mathcal{L} of the decomposition. Then, we can quantify how the pair (v_d, γ) provides a good fit to the kinematics of the j -th population as

$$L_j(v_d, \gamma) \propto \mathcal{L}(R_{e,j}, n_j, \dots) \times e^{-(\hat{\sigma}_j(v_d, \gamma) - \sigma_j)^2 / (2\delta\sigma_j^2)} \quad (21)$$

marginalized over the other parameters (R_e, n etc).

As a consequence, each *pair* of sub-populations yields the scale-free density profile in the range spanned between the two effective radii considered. Let us recall that the enclosed mass in this class of models is $M(r) \propto r^{3-\gamma}$. Then, from Fig. 4, we can expect a steeper power-law density profile (i.e. higher γ) between the reddest and intermediate GCs and a shallower one between the intermediate and bluest population. Fig. 5 shows the posterior distribution of the density exponent and velocity parameter v_d when different pairs of populations are considered. In order to make the comparison with measured velocity dispersions easier, we have displayed the distribution of $v_d/\sqrt{3}$. The density profile is steeper between the red and intermediate-colour GCs and it becomes shallower in the region probed by the blue component, as expected.

The scale-free results exhibit an interesting behaviour, which helps deepen our insight. The averaged velocity dispersion decreases or increases with R_e depending on whether γ is larger or smaller than 2. *In other words, a dynamically hotter sub-population has a larger effective radius if the potential is shallower than isothermal, and a smaller effective radius if the potential is steeper than isothermal.* This is nothing more than the dimensional scaling $\sigma \propto v_c(R_e)$ with the proper coefficients given by eq. (18). The power-law exponents in Fig. 5 agree with this picture. More precisely, $\sigma_r > \sigma_i$ corresponds to $\gamma > 2$ in the region spanned by the red and intermediate-colour populations; conversely, $\gamma < 2$ between the intermediate and blue GCs matches $\sigma_b > \sigma_i$.

4.3 Broken Power-law Density

The shallower behaviour of $\rho(r)$ at large radii could be a spurious effect, since a single power-law may not be an adequate description over the whole extent of the tracers. A similar phenomenon was studied in Agnello et al. (2013), where broken power-law densities

$$\rho(r) = \rho_0 \frac{(r/r_0)^{-\gamma}}{(1 + r^2/r_s^2)^{(3-\gamma)/2}} \quad (22)$$

were analysed by means of scale-free models. A broken power-law density introduces an explicit length-scale, so that there are now three independent parameters. If dynamical measurements extend out to the break radius r_s , a scale-free model $\rho \propto r^{-\gamma_g}$ will strongly bias the estimated exponent towards $\gamma_g = 2$ and produce a non-monotonic dependence of γ_g on the true exponent γ .

Since we have three sub-populations at our disposal, we can demand that they satisfy the virial theorem simultaneously and infer the likelihood distribution for all the three parameters, ρ_0, γ and r_s . Instead of ρ_0 , we will continue to work with v_d as defined in eq. (20). This is no longer the circular velocity at R_d (unless $r_s \gg R_d$), but it makes the comparison with scale-free models easier. Once again, eq. (18) yields three velocity dispersions $\hat{\sigma}_j$ as functions of the two photometric parameters ($R_{e,j}, n_j$) and of the three dynamical parameters (v_d, γ, r_s).

In this case, we can incorporate the dynamical model within the sub-population partition, by using the three $\hat{\sigma}_j$ instead of σ_j in eq. (8) for the likelihood \mathcal{L} . This, in turn, is now a function of the three dynamical parameters (v_d, γ, r_s), in addition to the colour ($\langle c \rangle = \langle g-i \rangle, \Delta c$) and photometric parameters (R_e, n) for each of the three populations (red, intermediate and blue) plus the two normalisations, $\Sigma_r(0) = \Sigma_r(0)/\Sigma_b(0)$ and $\Sigma_i(0) = \Sigma_i(0)/\Sigma_b(0)$.

The marginalised likelihood in the colour parameters and fractions ($\Sigma_r/b, \Sigma_i/b$) is the same as for the model-independent partition of Section 4.2. However, two differences arise in the correlations between effective radii and in the posteriors inferred on the velocity dispersions ($\hat{\sigma}_b, \hat{\sigma}_r, \hat{\sigma}_i$). First, the model-independent partition privileges a strict ordering $R_{e,b} > R_{e,i}$ on the effective radii, whereas this dynamical decomposition allows for the possibility of reverse ordering. In particular, the posterior probability $P(R_{e,i} > R_{e,b})$ is approximately 40%. Secondly, there is a mismatch between the optimal dynamical parameters and the peak of the marginalised posterior distribution. The likelihood \mathcal{L} is maximised for $\gamma = 1.52$, $v_d = 455\sqrt{3}$ and $r_s = 0.26$ degrees, and corresponding velocity dispersions $\hat{\sigma}_b = 344$, $\hat{\sigma}_r = 313$, $\hat{\sigma}_i = 287$ kms⁻¹, which are compatible with the findings of Section 4.2. On the other hand, the *marginalised* posterior distribution in $\hat{\sigma}_j$ peaks at $\hat{\sigma}_b = 332$, $\hat{\sigma}_r = 296$, $\hat{\sigma}_i = 318$ kms⁻¹. If the condition $R_{e,b} > R_{e,i}$ is enforced, there is no way of obtaining the right ordering in the velocity dispersions. In other words, the broken power-law model requires the intermediate-colour population to be more extended than the blue one, in order to reproduce the best-fitting velocity dispersions. Alternatively, if the blue component is broader than the intermediate one, the velocity dispersions are reduced (for the blue and red populations) or increased (for the intermediate population) with respect to the outcome of the model-independent partition.

Fig. 6 shows the marginalised posterior distribution in the three dynamical parameters (γ, σ, r_s). The triple (v_d, γ, r_s) is correlated with the photometric parameters, a fact that is reflected in the peculiar behaviour of the inferred velocity dispersions explained above.

4.4 Luminous and Dark Matter Decomposition

A more faithful description of the system may consist of a luminous and dark matter decomposition. The luminous component may contribute to the steeper inner density, whilst the DM density may indeed be shallower than r^{-2} . In particular, we can consider the following model:

$$\rho_{tot}(r) = \frac{M_\star}{2\pi^2 r_\star r^2 (1 + r^2/r_\star^2)} + \rho_{0,d} (r/r_{0,d})^{-\gamma_d} . \quad (23)$$

The luminous profile (proportional to M_\star) is similar to the commonly used Jaffe (1983) profile, but is more convenient computationally. The radius r_\star is chosen so that the effective radius is equal to that of the starlight, $R_{e,\star} = 0.02$ degrees (see Harris 2009, and references therein). The free parameters are the stellar mass M_\star , the DM exponent γ_d and the DM circular velocity at R_d (which we choose without loss

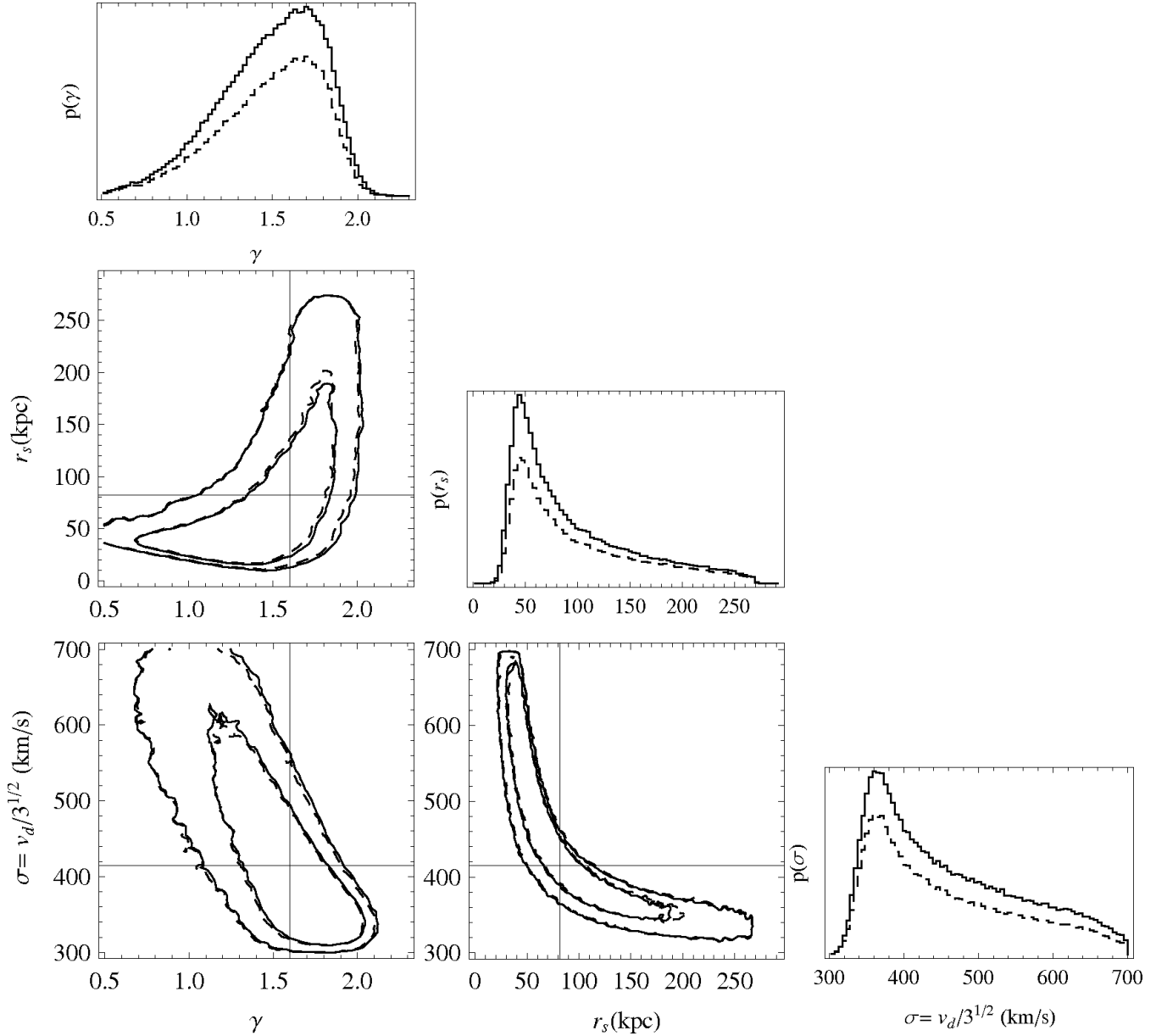


Figure 6. Posterior on the parameters of the broken-power-law model of eq (22). We show the 68% and 95% confidence regions of the marginalised posterior distributions in the space of power-law slope γ , logarithm of the scale radius $\log_{10}(r_s)$ (arcsec) and velocity dispersion $v_d/\sqrt{3}$ (in km/s), besides the resulting distributions of these three parameters. Full lines: contours from the full parameter space, including models with $R_{e,i} > R_{e,b}$. Dashed lines: contours from the marginalisation over parameter space with $R_{e,i} < R_{e,b}$. The inference on dynamical parameters is then independent of the morphological (R_e, n) ones. The axes in the correlation plots intersect at the best-fitting values (cf. Sect. 4.3), which do not always correspond to the marginalised ones.

of generality as 1000 arcsec):

$$v_d = \sqrt{\frac{4\pi G \rho_{0,d} R_d^{2-\gamma_d} r_{0,d}^{\gamma_d}}{3 - \gamma_d}}, \quad (24)$$

plus the photometric and colour parameters.

The virial theorem gives three velocity dispersions $\hat{\sigma}_{b,r,i}$ as a function of (M_*, v_d, γ_d) and the photometric parameters, which are used in the maximum likelihood method via eq. (8). Now the posterior distribution reproduces the same features of the model-independent partition, both in terms of effective radii ($R_{e,b} > R_{e,i}$) and velocity dispersions, which are listed in Table 5. Figure 7 shows the marginalised like-

lihood in terms of the dynamical parameters. Within this model, the posterior has a peak at $M_* = 5.5 \times 10^{11} M_\odot$, $\gamma_d = 1.6$ and $v_d = 230\sqrt{3}$. The 1σ confidence intervals for the parameters are: $1.31 < \gamma_d < 1.76$, $3.0 < M_*/(10^{11} M_\odot) < 6.9$, and $170 < v_d/\sqrt{3} \text{ km/s} < 280$. Then, the underlying profile is reliably described as the superposition of simple luminous and dark components, although the inferred dynamical parameters inherit the large uncertainty from the photometric parameters. At very low values of M_* , the solution with $\gamma_d = 2$ and $v_d = 350\sqrt{3} \approx 600 \text{ km/s}$ is available. This conspiracy between luminous and dark matter to produce an overall $\rho \propto r^{-2}$ profile has been observed

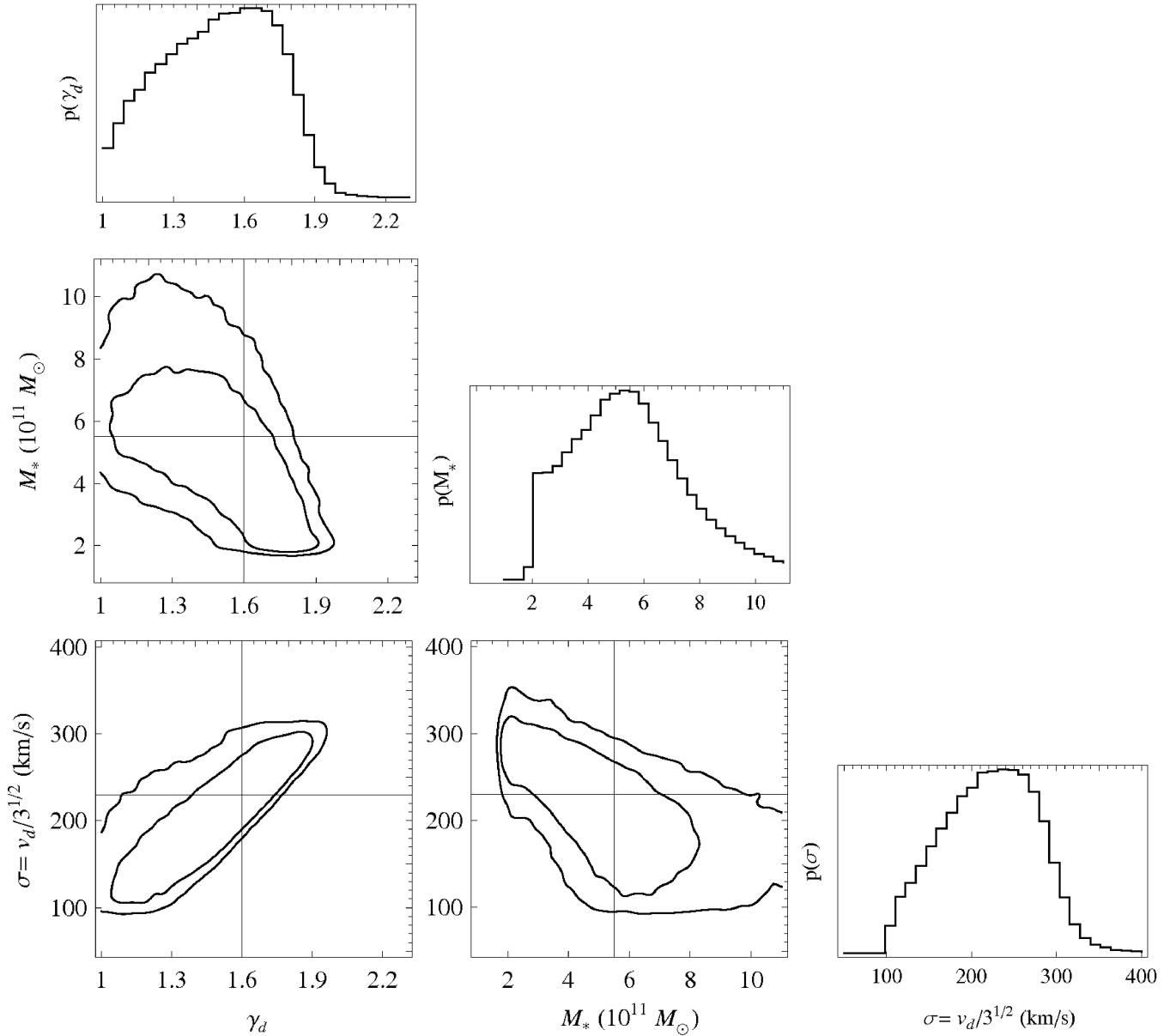


Figure 7. Posterior on the parameters in the luminous and dark decomposition given in eq (23). We show the 68% and 95% confidence regions of the marginalised posterior in the space of of stellar mass M_* (in units of $10^{11} M_\odot$), dark halo power-law slope γ_d , and velocity dispersion $v_d/\sqrt{3}$ (in km s^{-1}). The axes intersect at the best-fitting values, which roughly correspond to the peaks of the marginalised distributions.

in other cases (Treu & Koopmans 2004; Humphrey & Buote 2010; Remus et al. 2013),

Finally, it is interesting to compare the best-fitting dark matter masses to our earlier estimates in Table 4. Integrating our dark matter density law out to $0.46^\circ \approx 135$ kpc, the dark matter mass is $8.0 \times 10^{12} M_\odot$. Added to the luminous mass of $5.5 \times 10^{11} M_\odot$, gives a total mass of $8.6 \times 10^{12} M_\odot$, reassuringly close to the value of $9 \times 10^{12} M_\odot$ given by the cruder estimates of Sect. 4.1. The total mass at 135 kpc, inferred from the likelihood on the dynamical parameters, has sizeable uncertainties, namely $M_{\text{tot}}(135\text{kpc}) = 8.0^{+1.0}_{-4.0} \times 10^{13} M_\odot$.

4.5 Evidence for DM Contraction

The model with luminous and dark components as in eq. (23) predicts kinematic and photometric properties compatible with the model-independent results of Section 3.2. At the same time, it yields information on the best fitting power-law slope of the dark matter. Although the posterior distribution of γ_d in Fig. 7 is broad, the peak of the distribution lies at $\gamma_d \approx 1.6$. This suggests that the underlying DM density may be steeper than r^{-1} , which is predicted at small radii by cosmological DM-only simulations (Dubinski & Carlberg 1991; Navarro et al. 1996, hereafter NFW). However, this result may be the outcome of a model, where the DM halo has a scale-free density. Perhaps the DM density is more appropriately described by a broken power-

law similar to eq. (22), which is a simple generalisation of the NFW profile.

We can test whether this is a mere consequence of the adopted model by allowing the DM density to be a broken power-law, parameterized by an inner exponent γ_d , a normalisation v_d and a break-radius r_s as in eq.(22). Within the virial method, we can use these as free parameters in the decomposition into three populations, provided the luminous mass M_* is kept fixed. The procedure can be repeated for different fixed values of M_* , to quantify the change in the inferred γ_d , v_d and r_s . The posterior distribution (marginalised over colour and photometric parameters) has the same width regardless of the chosen value for M_* , whereas the maximum likelihood values vary as

$$\begin{aligned}\gamma_d(M_*) &\approx 1.575 - 0.075 \left(\frac{M_*}{10^{11} M_\odot} - 5.4 \right), \\ v_d/\sqrt{3} &\approx 297.5 - 12.5 \left(\frac{M_*}{10^{11} M_\odot} - 5.4 \right) \text{ km/s}, \\ r_s &\approx 590 - 67 \left(\frac{M_*}{10^{11} M_\odot} - 5.4 \right) \text{ arcsec}.\end{aligned}\quad (25)$$

The uncertainties are symmetric for γ_d and $v_d/\sqrt{3}$ and amount to 0.2 and 45 kms^{-1} respectively, whereas r_s has a very skew posterior distribution around the peak value. We can marginalise the inner exponent over M_* , by considering a sum of Gaussians with mean $\gamma_d(M_*)$ and dispersion $\delta\gamma = 0.2$ and weighing M_* with the virial likelihood (Section 4.4 and Fig. 7). The resulting exponent is $\overline{\gamma_d} = 1.55 \pm 0.25$, which is still appreciably different from the simple NFW prediction $\gamma_d = 1$. Furthermore, a decomposition relying strictly on a NFW halo gives inversions in the effective radii and velocities dispersions similarly to that found in Section 4.3. These findings lend decisive support in favour of DM contraction, with a DM density exponent $\gamma_d \approx 1.6$.

5 JEANS ANALYSIS

In the previous sections, the dynamics of the system has been explored through the projected virial theorem, which is derived through integration of the Jeans equations over configuration space.

If instead the Jeans equations are employed directly, then the sub-populations are described by the previous parameters ($R_{e,j}$, n_j , ...) and also by their anisotropy profiles

$$\beta(r) = 1 - \frac{\langle v_t^2 \rangle}{2\langle v_r^2 \rangle}. \quad (26)$$

Here, $\langle v_r^2 \rangle$ and $\langle v_t^2 \rangle$ are the radial and tangential velocity second moment, while r is the spherical polar radius.

In our analysis, we will use

$$\beta(r) = \frac{\beta_{\text{in}} r_a^2 + \beta_{\text{out}} r^2}{r_a^2 + r^2} \quad (27)$$

so that the inner anisotropy is β_{in} and the outer anisotropy β_{out} , whilst r_a is a transition radius. Here, we use eq (16) of Paper I to calculate the velocity dispersions averaged within radial annuli, and thus build kinematic profiles, which are then used to separate the populations similarly to the procedure followed in the previous Sections.

As already noted in Section 3.2, the use of parameterised kinematic profiles, such as in eq (9), does not improve the inference on average velocity dispersions, which are linked to the mass model via the virial theorem. Thus, we can expect that the Jeans analysis will provide information on the velocity anisotropy profile, but no significant additional constraints on the mass model.

5.1 Dynamical parameters

The results of the previous sections suggest that we restrict attention to the two-component mass-model of eq. (23), which is both realistic (Section 4.4) and reliable (Section 4.5).

Once the likelihood is marginalised over the colour and photometric parameters, the posterior in the dynamical parameters (M_* , v_d , γ_d) and the anisotropies is left. In Fig. 8, we show the inference on the inner anisotropies β_{in} of the three populations and on the dynamical parameters. The mass-anisotropy degeneracy is evident in the correlation between β_{in} and the DM parameters (γ , v_d), for each population. There is a weaker correlation between this set of parameters and the remaining ones (β_{out} , r_a). The posterior distribution of the dynamical parameters is similar to what we found in Section 4.4.

Compared to the virial results, the uncertainties on the DM parameters are slightly smaller but comparable. The likelihood profile is very shallow in M_* and the marginalised posterior inherits its profile from the behaviour of the likelihood in the other parameters. For example, at small values of M_* a wider range of anisotropies is allowed, giving a larger value for the integrated likelihood towards the low-mass end. This is another effect of the mass-anisotropy degeneracy. Having chosen a particular model for $\beta(r)$, Jeans methods will privilege a narrower range of dynamical parameters, whereas the virial results account for the whole generality of anisotropy models that could be chosen.

5.2 Anisotropies

The inferences on outer anisotropies and anisotropy radii are shown in Fig. 9. The value of r_a/R_e is not well constrained, the only reliable information being a lower limit $r_a \gtrsim R_e/2$ (cf Sonnenfeld et al. 2012, for a similar case). A narrower distribution is obtained if the anisotropy at the effective radius $\beta(R_e)$ is considered, which is shown in the last panel.

A glance at the inferred anisotropy profiles (figures 8 and 9) shows some interesting features of the system. The red GCs are on slightly tangential orbits, with nearly constant anisotropy. This may be simply consistent with the supposition that GCs at sufficiently small distances from M87 are tidally shredded at pericentric passage unless their orbits have sufficient angular momentum (cf Webb et al. 2013). This creates a *loss-cone* (Bahcall & Wolf 1976) in velocity space, whose aperture decreases with the distance from the center (as r^{-2} for a flat rotation curve). The intermediate-colour GCs, which orbit at larger radii, have a mildly radial anisotropy, $\beta(r) \approx 0.3$.

On the other hand, the blue GCs have approximately isotropic orbits at smaller radii and a mildly tangential velocity dispersion tensor at larger distances. This behaviour

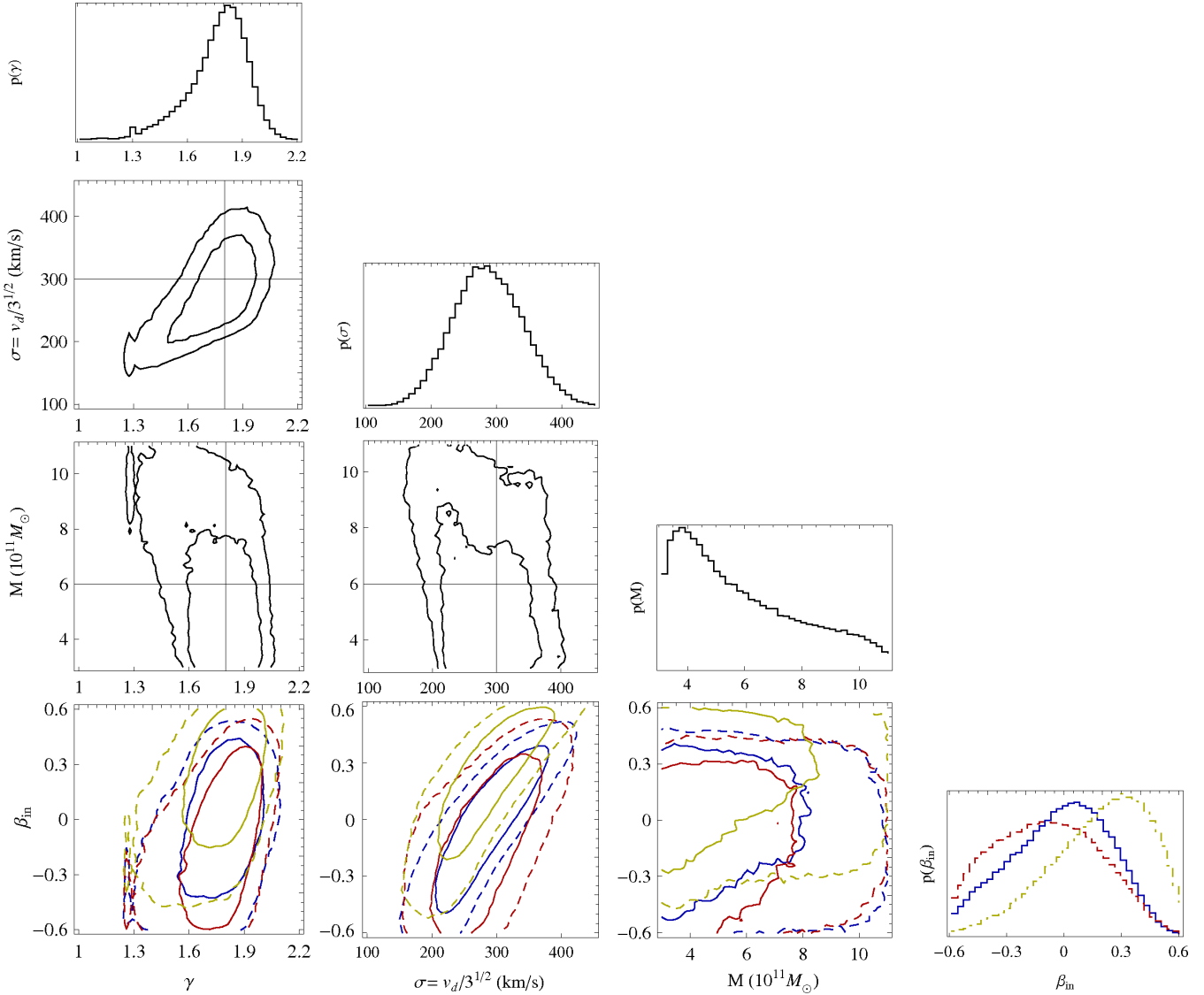


Figure 8. Posterior on the parameters in the luminous and dark decomposition, together with the inner anisotropy, as inferred from the Jeans equations (Section 5 and Paper I). We show the 68% and 95% confidence regions in parameter space, together with the marginalised posterior distribution on the single parameters. Different colours indicate the parameters of different populations: blue full lines (red dashed lines, and yellow dot-dashed lines) stand for the blue (red, and intermediate) populations. The axes intersect at the best-fitting values; the best-fitting M_* does not correspond to the peak of the marginalised likelihood, because of the mass-anisotropy degeneracy, whereas γ and σ follow the same covariance as in the virial analysis.

cannot be explained with arguments of tidal disruption, since the loss-cone in velocity space is already small enough for the intermediate-colour GCs to survive on mildly radial orbits. There is, however, another dynamical phenomenon that can produce such an anisotropy profile, which is related to the accretion process of the central object. If M87 has accreted mass on sufficiently slow time-scales, then radial orbits are dragged towards the center more efficiently than tangential ones (Goodman & Binney 1984; Lee & Goodman 1989; Cipollina & Bertin 1994), which then contributes to a net tangential anisotropy in the outer parts. This scenario of (approximately) adiabatic contraction would also agree with the enhanced DM exponent, $\gamma_d \approx 1.6$, significantly larger than unity (Sections 4.4, 4.5).

Orbital time-scales are shorter at smaller distances.

Then, if the intermediate-colour GCs are coeval with the blue ones, they must have experienced a similar phenomenon. The inferred positive value of β_{out} for this population suggests that they have been falling onto M87 on preferentially radial orbits, differently from the more distant and bluer GCs.

6 CONCLUSIONS

The giant elliptical galaxy M87 is surrounded by one of the largest known populations of globular clusters (GCs), exceeding 10^4 in total (Tamura et al. 2006). This extensive swarm of GCs offers a wealth of information regarding M87's mass profile, together with tantalizing evidence about mechanisms of its formation and evolution. We have been able to

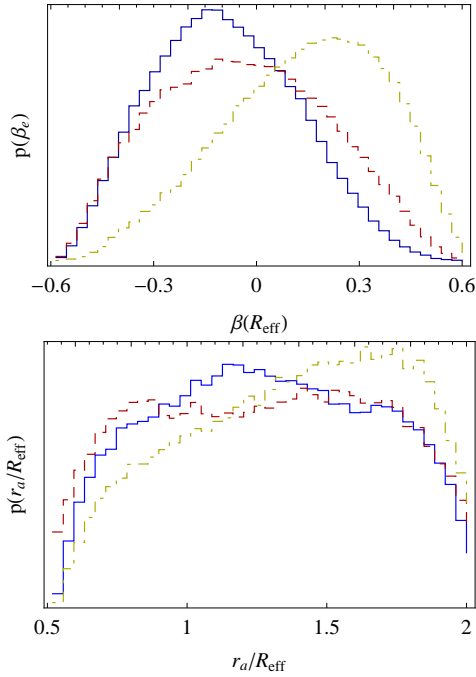


Figure 9. Marginalised posterior distribution on the anisotropy parameter at the effective radius $\beta_e = \beta(R_e)$ (top panel) and the ratio of anisotropy radius to the effective radius r_a/R_e (bottom panel). Colour coding as before: blue solid lines (red dashed lines, and yellow dot-dashed lines) stand for the blue (red, and intermediate) populations.

answer some questions, thanks to the availability of a new data set of positions, colours and velocities with the largest radial extent and best kinematic accuracy to date (Strader et al 2011, or S11).

6.1 Multiple populations

The colour distribution and kinematics of the GCs at different radii suggest a partition into multiple sub-populations (see, for example, Harris 2009, S11 and references therein). Here, we have shown that a separation into three components is statistically preferred over ones into two or four populations for M87’s GC system. To separate the three components (blue, intermediate and red GCs), we have exploited a maximum-likelihood method in which the information from position, colour and velocity is used jointly. The average velocity dispersion and the colour-distribution parameters of each population are robustly determined, whereas the Sérsic indices and half-mass radii are affected by sizeable uncertainties (see for example, Table 2). The uncertainty in turn is a consequence of the limited number ($N \approx 420$) of bona fide GCs with high quality spectroscopic data.

A case for more than two GC populations has been made also in studies of other galaxies, most notably NGC 4365 (Brodie et al. 2005). The same system has been reanalysed by Blom et al. (2012), where the colours of the three sub-populations have been linked to their different kinematics and viable formation scenarios. Here, we have shown how the dynamics (density profile and anisotropies) can be used, within a robust Bayesian analysis, to optimally decompose

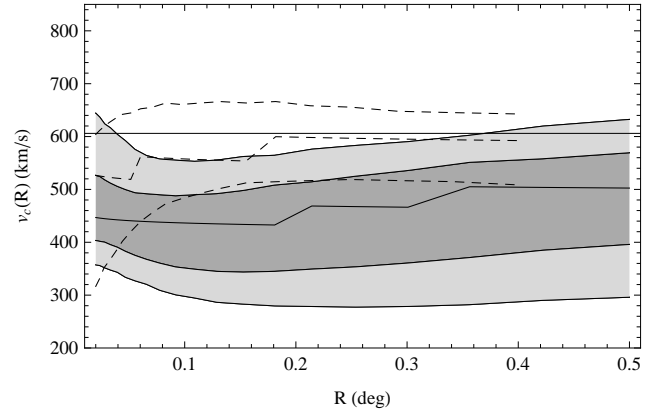


Figure 10. Circular velocity curve as inferred from M87’s GCs. The grey-shaded regions mark 68% and 95% confidence levels for the model with luminous and dark matter (Sect.4.4). The dashed lines show the average and 68% confidence level inferred from the dynamical model with a broken power law (Sect.4.3), which however is discarded after comparison with the model-independent partition of Sect.3.2 and is shown here just for the sake of completeness. The horizontal axis at $v_c = 606$ km/s corresponds to the case of a flat circular velocity curve, $\rho_{tot} \propto r^{-2}$.

the GC system in sub-populations as well as compare formation scenarios with observational properties.

6.2 Masses

The kinematics of the GCs are a direct probe of the total mass distribution of M87. Each sub-population with its own global velocity dispersion and photometric parameters has its own virial pinch radius (eq.15,16 and Paper I). Therefore, the red, intermediate and blue GC populations provide us with three enclosed masses at three different pinch radii (see Table 4). The mass estimate at the pinch radius of the red GCs is corroborated by the results of more elaborate techniques applied to the stellar and GC populations (Murphy et al. 2011). The method is robust, as it is derived as a blind average over different possible halo mass models. The largest uncertainty is in the determination of the effective radius, and so the covariance between pinch radius and enclosed mass follows the scaling $M(R_*) \propto R_*$ inherited from eq (15). This does not imply that the underlying mass distribution corresponds to that of a flat rotation curve. Indeed, there is a danger that the uncertainties scatter solutions along the locus of flat rotation curve models, which really amounts to fitting the covariance of the data rather than the underlying mass profile.

The mass profile can also be constrained by using dynamical modelling to obtain the velocity dispersions, which are then used to separate the three components. The first step consists of applying the Projected Virial Theorem (eq. 18). This has the advantage of involving the surface-density averaged (i.e. global) velocity dispersions only, thus avoiding systematic uncertainties arising from the mass-anisotropy degeneracy. By comparing the results of the separation relying on the virial theorem with those of the model-independent decomposition, we have been able to validate the luminous and dark matter model of eq (23), placing constraints on the luminous mass, dark matter (DM) density ex-

ponent and normalisation. The luminous and dark masses at different locations, together with their uncertainties, are discussed in Sections 4.1, 4.3 and 4.4. Figure 10 displays these results in terms of the inferred circular velocity at different locations.

The inferred density exponent is $\gamma \approx 1.6$, if a scale-free DM profile is used. When a broken power-law model for the DM halo is adopted, the inner exponent is still in the same range ($\gamma \approx 1.55$). This is consistent with the findings of Sonnenfeld et al. (2012) for an early-type galaxy of similar mass and can be interpreted as evidence for DM contraction. The scale-radius $r_s \approx 600$ arcsec where the DM density profile changes from $\sim r^{-\gamma}$ to $\sim r^{-3}$ is somewhat smaller than the effective radius of the intermediate-colour population, but appreciably larger (or smaller) than the effective radius of the red (or blue) component.

6.3 Comparison with X-ray Studies

Studies in the literature compare the masses measured from the hot X-ray gas to those inferred from GC kinematics (Murphy et al. 2011). The X-ray masses are typically lower than the ones from GCs at small radii and larger at large radii. The discrepancy at small distances has been alleviated by invoking some amount of non-thermal support for the gas, but the disagreement at large distances would still hold (Humphrey et al. 2013). It is fair to say that previous studies relied on velocity dispersions that were appreciably overestimated, so the role of non-thermal motions is far from clear. The complex and asymmetric X-ray photometry of the gas calls for further caution in deriving masses from the X-ray gas based on assumptions such as spherical symmetry.

Another worry is the assumption that the gas is exactly in hydrostatic balance. If $u(r)$ is the radial velocity of the gas, then from the Euler equation for a fluid flow with pressure p and density ρ

$$\partial_t u + u \frac{\partial u}{\partial r} \equiv \frac{du}{dt} = -\frac{1}{\rho} \partial_r p - \frac{GM(r)}{r^2}, \quad (28)$$

we see that depending on how the accreted gas settles towards hydrostatic equilibrium (i.e., positive or negative acceleration), the underlying mass profile may be under or over-estimated. This aspect has been discussed in detail for other systems (Ciotti & Pellegrini 2004; Pellegrini & Ciotti 2006), and there is evidence against hydrostatic equilibrium. Given the high virial mass of M87, it is plausible to assume that gas has been accreting onto it in the hot mode (Birnboim & Dekel 2003), thus slowly setting towards hydrostatic balance after shocking at the virial radius. If the gas has bounced off the central regions, as a consequence of pressure build-up from the converging inflow, then it might be still experiencing a negative acceleration (from positive to zero velocity) and the mass from X-ray studies may be an underestimate. If at large distances the gas is still passing from negative to zero velocity, then the convective acceleration is positive and the X-ray mass may be an overestimate.

Das et al. (2010) found that the X-ray velocity curve of M87 rises from 500 km/s at 6 kpc (the effective radius of the starlight) to 600 km/s at large radii. They suggest that these numbers are susceptible to a 10 % uncertainty due to systematic effects. This takes them to values close to ours

inferred from dynamical modelling, which range from 500 km/s at 6 kpc to 530 km/s at large radii. We conclude that there is tendency for the X-ray masses to be slightly over-estimated at large radii, but that the discrepancy is within the uncertainties.

6.4 Orbits

The results from the Jeans analysis suggest an interesting distribution in velocity space. On the one hand, the GCs at small radii are consistent with a loss-cone distribution function that privileges tangential orbits, as GCs on radial orbits suffer tidal disruption near pericentric passage. On the other hand, at large distances the anisotropy profile of the blue GCs tends again towards mild tangentiality, which cannot be explained by loss-cone arguments. A more plausible scenario would invoke accretion of external material onto M87 once the GCs are already in place. The build-up of mass has a different effect on GC orbits depending on their angular momenta. If accretion proceeds slowly enough, this gives rise to a tangentially biased velocity-dispersion tensor in the outer parts, together with a contraction of the DM density profile at smaller distances.

Inference on the structure in velocity space has been possible within the framework of the simple analysis presented here. This is based on the Jeans equations and the approximation of spherical symmetry, which is appropriate as a first step in the case of M87. The use of more refined modelling techniques, such as non-spherical Jeans modelling, orbit-based or made-to-measure, would be useful to encode additional effects like possible flattening of the dark halo and more elaborate velocity distributions. However, the analysis of Amorisco & Evans (2012) and the sizeable statistical uncertainties on the surface-density parameters suggest that more advanced techniques will have to wait until larger data sets are collected.

ACKNOWLEDGMENTS

We are indebted to Jay Strader for his observational effort on the M87 GC system, which has resulted in a sound basis for this work. We thank Luca Ciotti, Daniel Wang, Cathie Clarke and Andy Fabian for helpful discussions on X-ray masses, and Vasily Belokurov for significant feedback on the manuscript. We thank the anonymous referee for a very detailed report, which helped improve the manuscript considerably. AA acknowledges financial support from the Science and Technology Facilities Council (STFC) and the Isaac Newton Trust. AJR and JPB were supported by National Science Foundation grants AST-0909237 and AST-1211995.

REFERENCES

- Abadi, M. G., Navarro, J. F., Fardal, M., Babul, A., & Steinmetz, M. 2010, MNRAS, 407, 435
- Agnello, A., & Evans, N. W. 2012a, MNRAS, 422, 1767
- Agnello, A., & Evans, N. W. 2012b, ApJL, 754, L39
- Agnello, A., Auger, M. W., & Evans, N. W. 2013, MNRAS, 429, L35

- Agnello, A., Evans, N. W. 2013, MNRAS, submitted (Paper I)
- Amorisco, N. C., & Evans, N. W., 2011, MNRAS, 411, 2118
- Amorisco, N. C., & Evans, N. W. 2012, MNRAS, 424, 1899
- Amorisco, N. C., Agnello, A., & Evans, N. W. 2013, MNRAS, 429, L89
- Bahcall, J. N., & Wolf, R. A. 1976, ApJ, 209, 214
- Bird, S., Harris, W.E., Blakeslee, J.P., Flynn, C. 2010, AA, 524, A71
- Birnbom Y., Dekel, A. 2003. MNRAS, 345, 349
- Blom, C., Forbes, D. A., Brodie, J. P., et al. 2012, MNRAS, 426, 1959
- Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27
- Brodie, J. P., Strader, J., Denicoló, G., et al. 2005, AJ, 129, 2643
- Brodie, J. P., & Strader, J. 2006, ARA&A, 44, 193
- Brodie, J. P., Romanowsky, A. J., Strader, J., & Forbes, D. A. 2011, AJ, 142, 199
- Burnham, K. P., Anderson, D. R. 2002, Model Selection and Multimodel Inference; Springer, 2nd ed.
- Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
- Churazov, E., Tremaine, S., Forman, W., et al. 2010, MNRAS, 404, 1165
- Ciotti, L., & Bertin, G. 1999, AA, 352, 447
- Ciotti, L., & Pellegrini, S. 2004, MNRAS, 350, 609
- Pellegrini, S., & Ciotti, L. 2006, MNRAS, 370, 1797
- Cipolina, M., Bertin, G., 1994, AA, 288, 43
- Coccato, L., Gerhard, O., Arnaboldi, M., et al. 2009, MNRAS, 394, 1249
- Cohen, J. G., & Ryzhov, A. 1997, ApJ, 486, 230
- Cohen, J. G. 2000, AJ, 119, 162
- Das, P., Gerhard, O., Churazov, E., & Zhuravleva, I. 2010, MNRAS, 409, 1362
- Deason, A. J., Belokurov, V., Evans, N. W., & McCarthy, I. G. 2012, ApJ, 748, 2
- Dekel, A., Stoehr, F., Mamon, G. A., et al. 2005, Nat, 437, 707
- Douglas, N. G., Arnaboldi, M., Freeman, K. C., et al. 2002, PASP, 114, 1234
- Dubinski, J., & Carlberg, R. G. 1991, ApJ, 378, 496
- Goodman, J., & Binney, J.J. 1984, MNRAS, 207, 511
- Hanes, D. A., Côté, P., Bridges, T. J., et al. 2001, ApJ, 559, 812
- Harris, W. E. 2009, ApJ, 703, 939
- Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924
- Huchra, J., & Brodie, J. 1987, AJ, 93, 779
- Humphrey, P.J., Buote, D. 2010, MNRAS, 403, 2143
- Humphrey, P. J., Buote, D. A., Brighenti, F., Gebhardt, K., & Mathews, W. G. 2013, MNRAS, 430, 1516
- Jaffe, W. 1983, MNRAS, 202, 995
- Janowiecki, S., Mihos, J. C., Harding, P., et al. 2010, ApJ, 715, 972
- Jeffreys, H., 1961, The Theory of Probability, Oxford University Press
- Johansson, P. H., Naab, T., & Ostriker, J. P. 2012, ApJ, 754, 115
- Jorgensen, I., Franx, M., & Kjaergaard, P. 1995, MNRAS, 273, 1097
- Kendall, M., & Stuart, A. 1979, London: Griffin, 1979, 4th ed.
- Lee, M. H., Goodman, J. 1989, ApJ, 343, 594
- Madrid J.P., Harris W.E., Blakeslee J.P., Gomez, M. 2009, ApJ, 7905, 237
- McLaughlin, D. E. 1999, ApJL, 512, L9
- Morganti, L., Gerhard, O., Coccato, L., Martinez-Valpuesta, I., & Arnaboldi, M. 2013, MNRAS, 431, 3570
- Mould, J. R., Oke, J. B., & Nemec, J. M. 1987, AJ, 93, 53
- Mould, J. R., Oke, J. B., de Zeeuw, P. T., & Nemec, J. M. 1990, AJ, 99, 1823
- Murphy, J. D., Gebhardt, K., & Adams, J. J. 2011, ApJ, 729, 129
- Napolitano, N. R., Romanowsky, A. J., Coccato, L., et al. 2009, MNRAS, 393, 329
- Navarro, J., Frenk, C.S., White, S.D.M. 1996, ApJ, 462, 563
- Nipoti, C., Treu, T., Leauthaud, A., et al. 2012, MNRAS, 422, 1714
- Pota, V., Forbes, D. A., Romanowsky, A. J., et al. 2013, MNRAS, 428, 389
- Remus, R.-S., Burkert, A., Dolag, K., et al. 2013, ApJ, 766, 71
- Richardson, T., & Fairbairn, M. 2013, arXiv:1305.0670
- Romanowsky, A. J., & Kochanek, C. S. 2001, ApJ, 553, 722
- Romanowsky, A. J., Douglas, N. G., Arnaboldi, M., et al. 2003, Science, 301, 1696
- Romanowsky, A. J., Strader, J., Brodie, J. P., et al. 2012, ApJ, 748, 29
- Schwarzschild, M. 1979, ApJ, 232, 236
- Schuberth, Y., Richtler, T., Hilker, M., et al. 2010, AA, 513, A52
- Schuberth, Y., Richtler, T., Hilker, M., et al. 2012, AA, 544, A115
- Sonnenfeld, A., Treu, T., Gavazzi, R., Marshall, P.J., Auger, M.W., Suyu, S.H., Koopmans, L.V.E., Bolton, A.S. 2012, ApJ, 752, 163
- Strader, J., Romanowsky, A. J., Brodie, J. P., et al. 2011, ApJS, 197, 33
- Stiavelli, M., & Bertin, G. 1987, MNRAS, 229, 61
- Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223
- Tamura, N., Sharples, R. M., Arimoto, N., et al. 2006, MNRAS, 373, 588
- Tortora, C., Napolitano, N. R., Romanowsky, A. J., Cappacioli, M., & Covone, G. 2009, MNRAS, 396, 1132
- Treu, T., & Koopmans, L. V. E. 2004, ApJ, 611, 739
- Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20
- Watkins, L. L., Evans, N. W., & An, J. H. 2010, MNRAS, 406, 264
- Webb, J. J., Harris, W. E., Sills, A., & Hurley, J. R. 2013, ApJ, 764, 124
- Wu, X., & Tremaine, S. 2006, ApJ, 643, 210
- Zubovas, K., & King, A. R. 2012, MNRAS, 426, 2751